Offsite Analysis for Run II

the Regional Analysis Center concept

plan for this discussion

1. the physics goals
2. the analysis goals
3. incorporating the World

Chip Brock

for DØRACE working group

The Question:

After an enormous investment in $, people, and time by the US

and for the first time in DØ, an enormous investment from European, South American, and Asian collaborating nations:

how do we get the most physics out of Run II?

The short answer is:

- by making use of all available computing/caching resources

  in a manner which

- maximizes the full intellectual potential of hundreds of DØ physicists

At first glance, this seems like an obvious sentiment...

the job ahead is sobering

Some jargon and normalization:

- our currency for characterizing the accumulation of data is the *time integrated luminosity* delivered by the accelerator and subsequently recorded by the apparatus.
- The units we use make use of the simple relationship

number of events of type *j*

$$N_j = \sigma_j \cdot \mathcal{L}$$

integrated luminosity

the cross section for event type *j*

- cross section: units of cm^2, or barns (=10^{-24} cm^2)
  - typical Tevatron cross sections are picobarns
  - Time-integrated luminosities are measured in cm^{-2}, or usefully: pb^{-1}, or now fb^{-1}

In Run I, we collected approximately 100 pb^{-1} of data

In Run IIa, we anticipate 2-3 fb^{-1}, and Run IIb...15-20 fb^{-1}

- This will be a staggering sample: ultimately 10PB of raw and derived data.

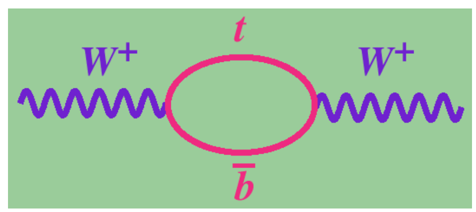
Discovery & Precision

Discovery measurements are inherently rare processes

- small □, requiring large $\mathcal{L}$
- In Run I, we required 100pb^{-1} to discover the top quark with ~40 events

Clue: quantum field theory allows the characterization of many Discovery channels through high-order radiative corrections

for example, calculation of



correlates M_W with m_{top}

a no-lose theorem!

- Precision measurements of parameters can usually be used to constrain the measurable characterizing a Discovery target
- We knew that **either** the top quark existed and m_{top} ~ 170 GeV/c^2 (a Discovery) **or** it didn't exist (which would have also been a major Discovery)

Precision measurements must be incredibly precise in order to make use of these correlations

- So, again, large $\mathcal{L}$

an important example

Standard Model hints at mass-generation mechanism

– Its realization is a spin-0 excitation called the Higgs Boson
– like top, its mass – *its existence* – is constrained by field theory

Two important Fermilab Measurables are m_{top} and M_W

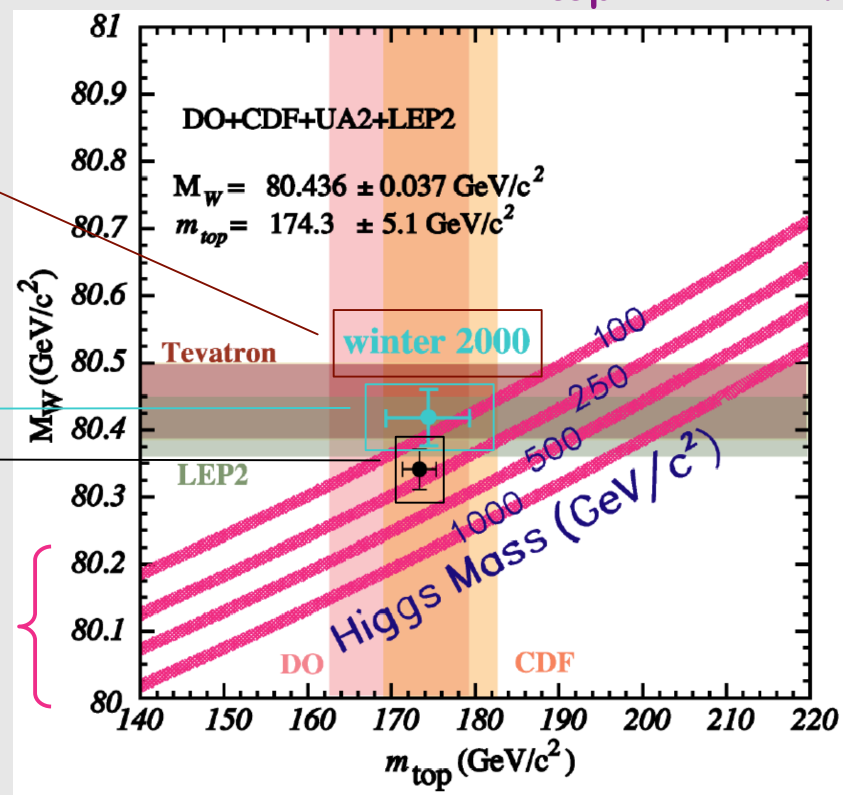
SM field theory correlates them:

hasn't changed in a while - a measure of our frustration!!

combination of FNAL-LEP

Run II uncertainties plotted at the M_W central value *before* Run I

each curve is a field theoretic calculation of the m_t - M_W constraint for different m_H



another no-lose theorem!

must constrain this plot with, for example,

determination of the mass of the *W* boson, esp systematic error

- It's mass is approximately 80 GeV/c^2
- In order to be interesting, in Run I δM_W had to be ~ 0.1 GeV/c^2 or ±0.1%

  It was – requiring 20,000 *W* bosons collected and:

  - ~5 years of effort by dozens of physicists
  - an understanding of the uniformity and calibration of the detector to a fraction of a percent – 2 years of effort alone
  - event simulations spanning literally hundreds of different theory and experimental parameter choices, each with samples of 20M…a day of processing for each

In Run II, to be consistent with the statistical precision:

- δM_W must be ~ 0.020-0.030 GeV/c^2 or ±0.025%

  enormous stress on detector understanding

  enormous burden in simulation

  requires experienced people as well as hardware

- ALSO, must measure δm_{top} to ±1-2 GeV/c^2 out of 175 GeV/c^2, or ±0.9%

Along with the correlated direct Discovery search for Higgs:

These are only three out of literally hundreds of important measurements.

Success at this level of physics with only capability *on site* @FNAL will require:

- computing capability beyond that which will be available
- more brains than the number of seats which will be available
- impossible network connectivity, for full collaboration engagement

an alternative

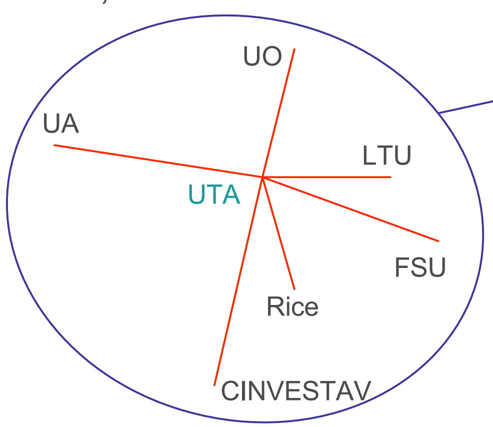
we are investigating tiered centers with goals to:

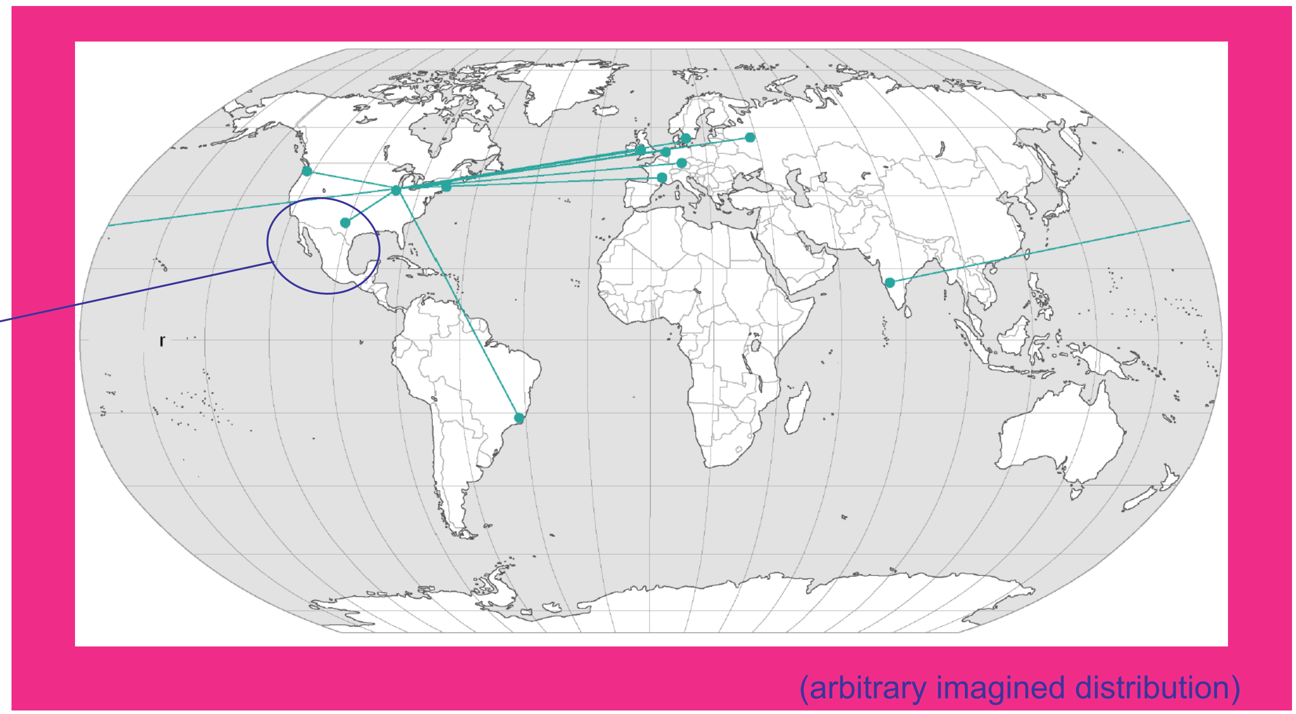
- distribute the data
- augment the processing FNAL cpu capabilities

allow countries to spend money “at home”

- engage the full collaboration: *especially engage senior experience*
- reduce the single-point traffic: *few high-bandwidth links to FNAL*

indeed, UTA has resources:





what is the concept?

Regional Analysis Center (RAC)

is an off-site facility that serves as a hub to nearby

Institutional Analysis Centers (IAC's)

There is a document which is a first look at the issues with some proposals for consideration:

DØ Note 3984: "Proposal for DØ Regional Analysis Centers"

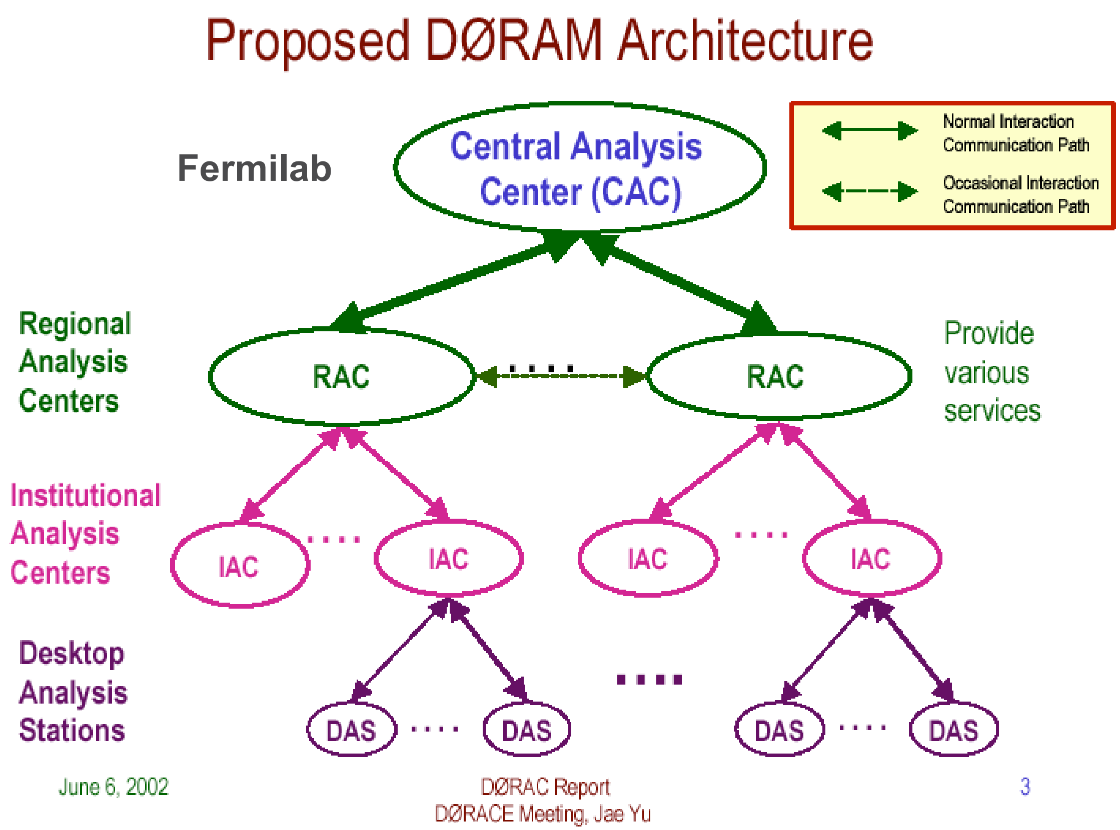
I. Bertram, R. Brock, F. Filthaut, L. Lueking, P. Mattig, M. Narain , P. Lebrun, B. Thooris , J. Yu, C. Zeitnitz

This discussion follows that paper and the consensus that led to it.

the RAC universe

the model

regional clusters of analysis institutions served by central, capable center...Regional Analysis Center (RAC)



I can get to 10-11:

- one each in Great Britain, Germany, France, Russia, and the Netherlands
- one in South America
- one in Asia (including India)
- one at Fermilab (CluDØ/CluB?)
- 3 more in the US - East, West, South

presumably distributed according to geographical, political, and/or infrastructure criteria...

why?

Because of the promise and complexity of the Run II data.

The worldwide investment in this run demands that we get all of the physics out

The old way…where the action's @FNAL:

- Everything done at FNAL, outside institutions stationed as many people in Illinois as affordable…and the faculty traveled
- Fermilab absorbed the cost of processing and data storage

The RAC way…where the action's @everywhere:

- Off-site institutions become full intellectual contributors:

  success in this difficult analysis will be guaranteed
- The job ahead to extract the physics demands it

  the resources required are available, but yet locally managed and owned

My opinion: health of HEP at US campuses (maybe everywhere?) needs @home-presence.

what about... THE GRID?

I know what you're thinking: "Is this the famous GRID?"

For DØ, perhaps not in its full glory – in our analysis lifetime

Can the enthusiasm of worldwide GRID proponents be justified? We'll see.

- **BUT** - some increasingly capable toolkit for resource balancing, job submission, data transfer, scheduling, statistics, metadata access, etc. *will come*, incrementally

We'll always be somewhere between no Grid and full Grid

- **BUT** our experiences will **certainly productively feed back into both LHC GRID planning and even generic GRID research and development**

We have a distributed data management tool now: SAM

- **SO** "GRIDifying SAM" or "SAMifying the GRID" is a major priority

With or w/out GRID, coordinating humans will be key

Need flexibility, replace humans with tools when stable and useful

Premium on stability, the analysis is not a GRID beta test facility

imagined RAC services

- **enhanced batch processing for region**
  IAC processing privileges at local RAC
- **data cache and delivery for region**
  RACs deliver not just to local IACs, but everywhere
- **database access for region**
  hopefully can rely on db proxies
- **data reprocessing for collaboration**
- **Monte Carlo production or service to local MC IAC sites**

notice what's not there:

ab initio reconstruction of raw data

– presume that the processing farm at FNAL will always keep up

code distribution

– we have a good code distribution directly from FNAL
  - it looks like it will scale
  - there might be a need for local support structures to triage questions/problems before they get back to FNAL
  - growing world-wide expertise with code dist., SAM, databases, etc.

what constitutes an RAC site?

- **I'll describe a minimum possibility which should satisfy our goals:**

  and scale it up to a "best' version

- **The bottom line for the system of RACs:**

  the totality of RAC's would have to be capable of:

  - reprocessing the data if required
  - complementing, not just replicating the FNAL storage capacity
  - significantly increasing the intellectual input to the whole analysis

RAC minimum requirements: ***geographic***

- **location, location, location**

  They have to be positioned in order to serve

  - Anticipate a few RACs - not more than ~10
  - try to distribute according to density of users
  - there will be other overriding considerations:

    network capabilities, political issues, language, funding, national goals, etc.

- **Networking**

  required: high-bandwidth, RACs to FNAL required ~ OC12c?

  required: high-bandwidth, RAC to local IACs

  recommended: high-bandwidth, RAC to all other RACs

RAC minimum requirements: ***computing***

- ## data storage

Generally thought desirable:

- **all DSTs on disk at the sum of all RAC's** -distributed randomly
  - *qualitatively different from FNAL complimentary*

    there, DSTs only available on tape
  - *hopefully the source for most reprocessing needs*
- **nearly all TMB files on disk at all RACs**
- **other formats on disk** (MC needs may involve local, high-capacity caching)
  - *derived formats*
  - *MC caching*
  - *database & SAM cache*
  - *temporary working cache ~10% of total*

results in ~30TB disk storage/year per minimal RAC for Run IIa

- ## cpu capability

Guess □ 5% x recent DØ planning model

- approximately 25 Linux nodes/year per minimum RAC for Run IIa

RAC minimum requirements: ***databases***

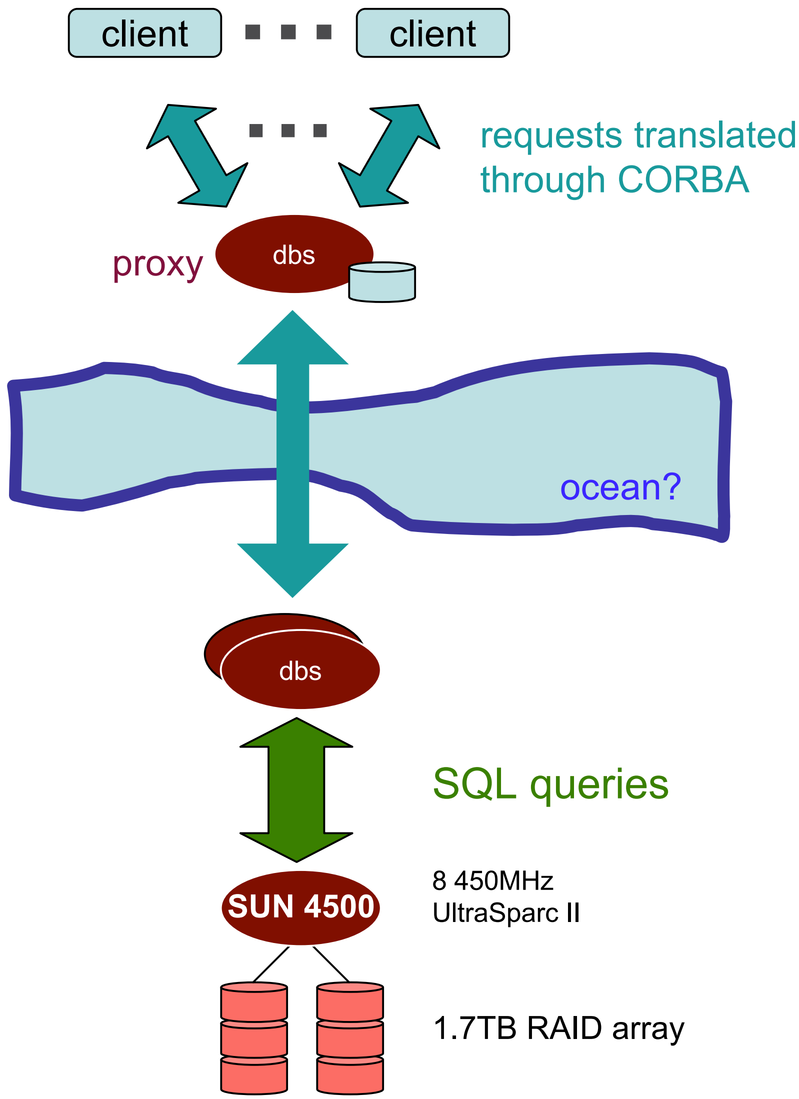
- **Oracle database access**

  presume implementation of proxied database servers

  - a feature of the current server upgrade

    every RAC would house a proxy server

  - this will hopefully be tested within the year



reprocessing: is this a requirement?

most think that it must be

Could be a major headache, or worse

- It happened a few times in Run I
- FNAL reconstruction farm will be fully busy with data coming in
- What would we do if we discovered a problem that required reprocessing from raw data?
  - study it hard...certainly no decisions over night
  - could decide to spend $ and triple the size of the farm and just do it *in situ*
  - or, could decide to use the set of RACs

    but, that involves getting raw data to them - not sure about this

DST format may handle most reprocessing needs

makes DST's availability on disk important

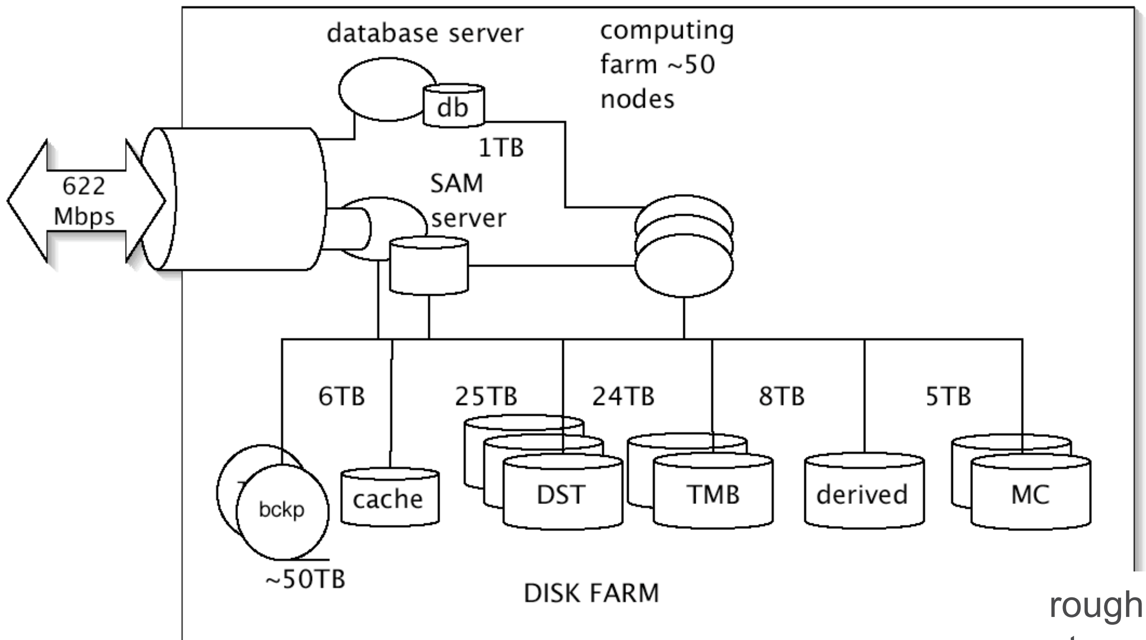
- reprocessing from DST would then be relatively straightforward at RAC's

makes the design of the DST a very important exercise

summary of the *minimum* RAC

For all of Run IIa

estimate something like this:



roughly 60TB of disk storage

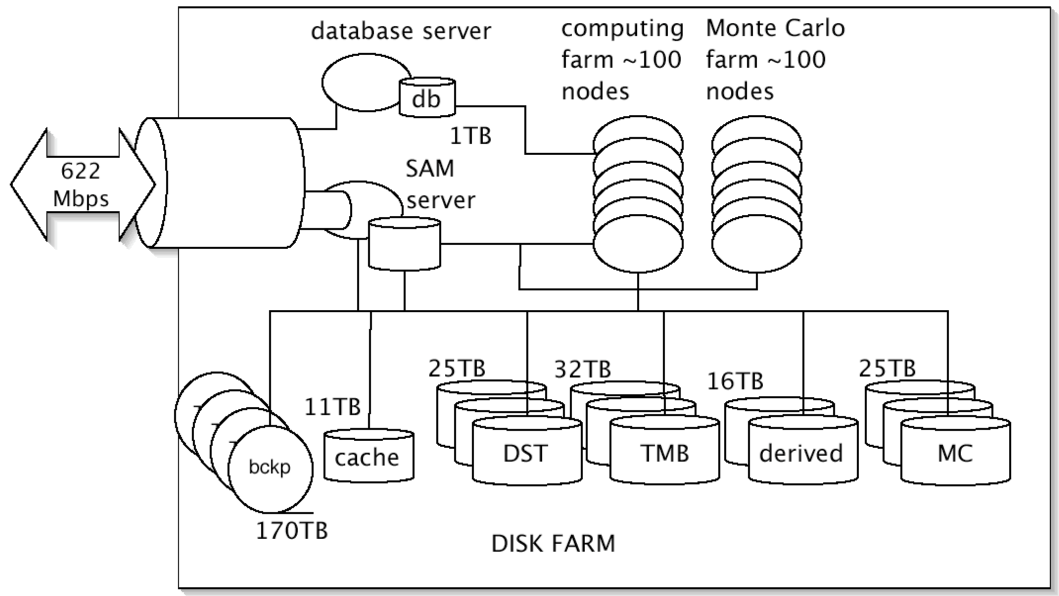
- this alone adds > 500 cpu's, deployed in a more efficient way - where the physicists are
- IAC's will also have additional capability
- All in host countries.

scale it up…to the Best RAC

Keep: DST storage as a common resource, SAM, db proxy server

Add:

- More batch computing: +50 nodes
- MC generation: + ~ 100 nodes
- More MC storage: +20TB
- All TMB: +8TB
- More derived data cache: +8TB
- More temporary cache: +5TB



So:

~100TB of disk and ~200 cpu's

a very serious system

desirable to have a few

may fit as parts of larger facilities

- No longer manageable by a single university department

still more requirements

Need support - *Best RAC* would require serious professional help

Not a simple setups - requires committed system management

Sharing with other experiments (not just DØ...not just HEP...or not even physics!) - is <u>inevitable</u>.

This probably has good aspects and not so good aspects

- loan of resources if crisis? funding? collaborative GRID R&D?
- But competing for resources and living with the politics of the GRID business might be frustrating.

Need a serious MOU structure

1 RAC dropping the ball freezes out the IACs <u>and</u> affects all of DØ

Need a worldwide coordination structure to keep the whole thing moving toward results.

– Will keep spokespeople & physics convenors awake at night...

we had a suggestion:

We thought of commissioning a *Prototype RAC Project*

- Identify, hopefully, a European institute (RAC_1) and a set of committed regional institutes (proto-IACs)
- Three goals:
  - TMBs are shipped in real time, continuously
  - and used by the proto-IACs to do physics
  - do it by winter?
- Declare success
  - Autopsy the effort and do it better the next time

    I suspect: sociology and management will present as big a set of complications as technology

    we need to understand this

volunteers are currently being solicited

I have some concerns

- **Video Conferencing**

  This is being worked on…will likely cost a moderate amount.

- **Culture**

  The collaboration and the Community need to buy into the idea that it's okay to be a post doc or student and live off-site.

- **At the risk of repeating myself:**

  The coordination of a world-wide analysis done in this fashion would be unlike anything HEP has tried to do before

  - Again, an important testbed for the advertised LHC plans

- **Review our few international computing projects**
- **The US funding agencies need to recognize it**

  LHC needs to embrace such an effort

- **Internal governance: how do we corporately decide?**
- **The document has 28 conclusions for consideration**

conclusion

- **This is early days -**
  - the idea was floated with the Run II Computing external review:

    "Both collaborations [should] develop more detailed plans for the coordinated use of remote computing facilities..."

    "The Committee congratulates the [sic] Dzero ... on its aggressive strategy to develop Regional Analysis Centers that would provide centralized regional access to data analysis resources. CDF's effort has been more modest..."
  - <u>we need to decide whether to do it and to what extent</u>

    we need a assessment of interest
  - we need to explore cooperation with LHC/NSF/DOE
- **Fundamental conclusion of our preliminary investigation:**

The analysis of Run II will be very difficult if it is fully collaborative, worldwide.

But, the analysis of Run II may not reach its potential if it is *not* fully collaborative, worldwide.

We have an obligation to insure physics success.

the rest of this is just storage of previous slides.

decisions required: 0. - 5.

12.1. Summary of Conclusions

Conclusion 0. Remote analysis capability with full access to the data, code, and collaborative analysis is necessary in order to satisfy the physics goals of Run IIa and IIb. A structured environment which systematizes and standardizes these services is the best way to implement this program.

Conclusion 1. It is anticipated that the FNAL processing farm will be sufficient for all of Run II primary reconstruction needs. RAC's are not envisioned for *ab initio* event reconstruction.

Conclusion 2. RAC-centered resource management is an important goal. While initially resource management may require considerable human organization, it is desirable to augment and replace that intervention with emergent GRID tools. The priorities assigned to tool deployment remains to be worked out with sufficient Use Case analyses and some real-world experience. Accordingly, the actual capabilities of the evolving system need to be carefully planned, biased toward smooth running rather than alpha or beta testing of GRID sites.

Conclusion 3. Continued evaluation of the number of off-site potential users and their anticipated needs should be undertaken very soon. A preliminary census has been done. The follow-up should include more detailed scenarios and/or capabilities for a more realistic assessment.

Conclusion 4. A complete review should be done of the planned data tiers with special attention paid to potential off-site reconstruction opportunities with DST's and analysis opportunities with a TMB's. This should be done before deploying the DST/TMB files.

Conclusion 5. Generally, RAC's need not be the sole sites of code distribution to their IAC's. Rather, at least for the early days, individual installation and updating can be done directly from the Fermilab site.

decisions required: 6. - 11.

Conclusion 6. Robust versioning and a scheme for guiding or automatically initiating stale file and directory deletions should be designed as soon as possible.

Conclusion 7. It may be important to precisely assess the degree of database access required for Monte Carlo production capability which includes overlaid events.

Conclusion 8. Early generation tools for interrogating the MC farm sites for available capability and tools for referring the actual job submission to those waiting sites are required now. Evolution of this system into a single step with full GRID deployment should be envisioned when it becomes available.

Conclusion 9. MC generation at RAC and/or IAC sites will be necessary in an increasingly wider scale as systematic uncertainties become a focus of measurements. Sufficient bookkeeping capabilities with the flexibility to imbed and modify the MC generation details will reduce the re-generation of already existing scenarios.

Conclusion 10. For the meantime, sufficient batch processing capability should be at each RAC to serve the needs of only the regional IAC's. What is uncertain is the amount of processing that this will entail and an effort to quantify this should be undertaken. Estimates will be made below.

Conclusion 11. In general, it seems reasonable to presume that in view of the limited computing resources available at FNAL, and of the improvements that are likely to be made to the algorithms used at present for the reconstruction of the collider data, some measure of reprocessing is expected to be an essential ingredient of the RAC's.

decisions required: 12. - 15.

Conclusion 12. Answers to questions 1-7 need to be in hand before reaching a conclusion about how to fully characterize generic RAC's.

1. How can the DST be ensured of serving as a useful basis for reprocessing? This is a timely design issue.
2. How will DST's be distributed for this reprocessing?
3. What strategies for raw data-level reprocessing can be designed?
4. If such strategies require RAC participation, how will the raw data make their way to the RAC's?
5. For any reprocessing operation, for the first time original data will reside at a location (many locations!) away from Fermilab. Do these derived data sets get transferred back to the central Fermilab facility? Certainly, the answer is "yes" for the TMB, by design.
6. Obviously, such a scheme involves significant networking and bookkeeping resources at the RAC's and the affordability of this should be understood.
7. How can the consistency of the reprocessing be guaranteed? This is not a matter only of code distribution, but also of hardware, OS, etc.

Conclusion 13. Reasonable database services could be achieved with the development of proxy database servers at each RAC.

Conclusion 14. Full testing of the performance of the new DBS implementation should be performed at the soonest available time.

Conclusion 15. A test installation of the proxy server idea at a remote site should be done in the near term.

decisions required: 16. - 22.

Conclusion 16. An evaluation of networking needs for remote analysis should be done for FNAL, U.S. RAC's, and overseas sites.

Conclusion 17. A robust and reliable SAM system at every worldwide site is essential. This means that adequate support for both development and operations must be provided.

Conclusion 18. All TMB records should be disk resident at all RAC sites twice. Total for Run IIa for TMB storage of 16TB disk per RAC.

Conclusion 19. Significant storage for project formats should be available at RAC's of the same order as the total TMB cache. Total for Run IIa for DERIVED storage of 16TB disk per RAC.

Conclusion 20. Complete DST data formats should be disk resident within the sum of the RAC sites: Total for Run IIa for DST data storage of 24TB disk per RAC, which presumes 10% of the total at each RAC site.

Conclusion 21. MC storage for per-demand generated events is primarily limited to the ROOTuple needs, with a nominal MCDST compliment as well: Total for Run IIa MC ROOTuple data storage of ~5TB disk per RAC and 10TB tape per RAC; MCDST disk storage of approximately 50TB, presumes 5% of the dataset on disk.

Conclusion 22. Temporary storage needs for staging of data and Monte Carlo analysis and ROOTuple generation are estimated to be 10% of the total of each data format: Total for Run IIa for temporary cache of ~11TB disk per RAC. A more accurate estimation of this need is required.

decisions required: 23. - 27.

Conclusion 23. Storage needs for serving db setups at RAC will not likely exceed a few TB for all data taking. Total for Run IIa for database/SAM needs of 1TB.

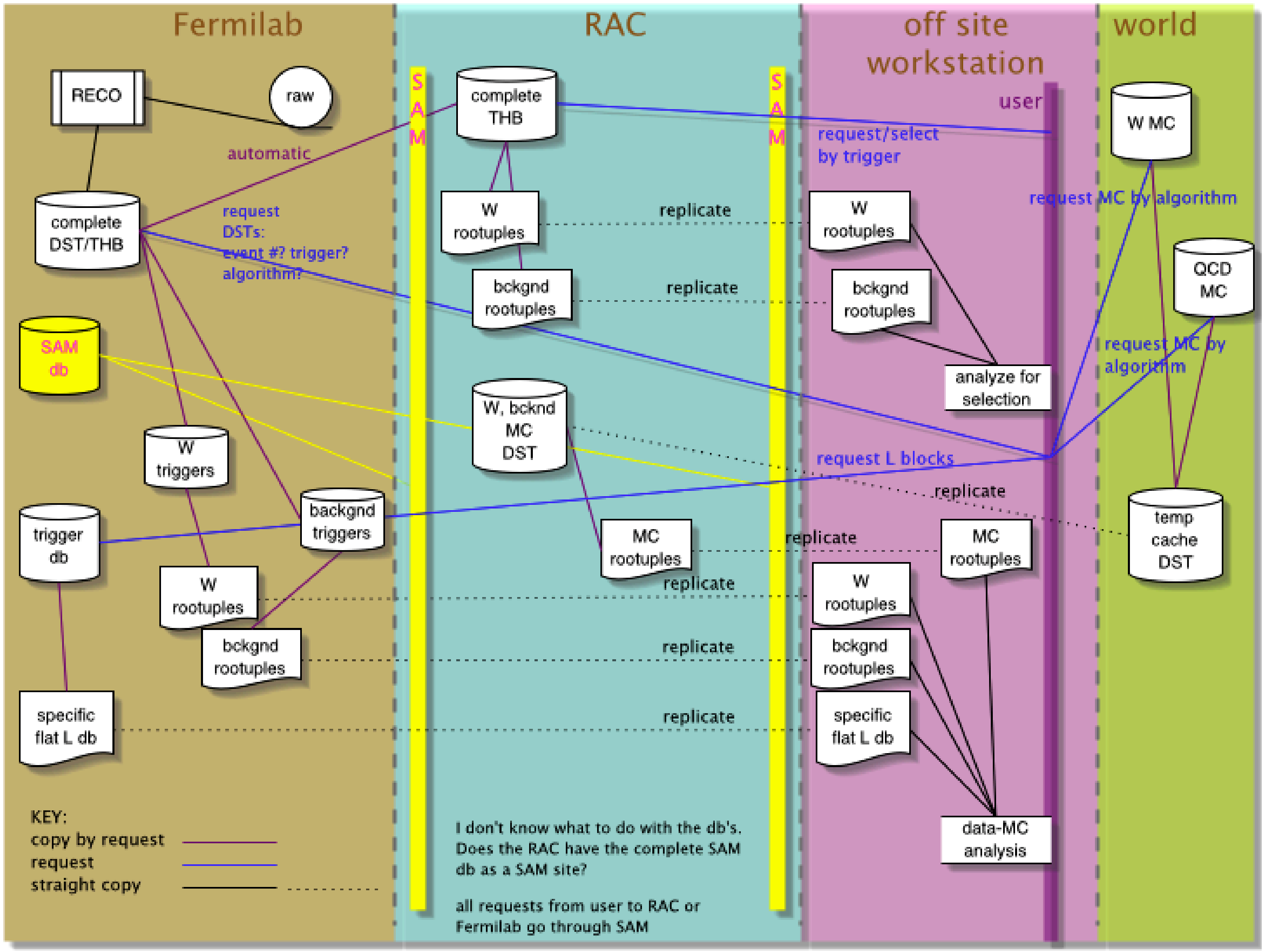
Conclusion 24. Batch resources per Category B RAC should be on the order of 50 nodes per year of modern PC nodes running Linux.

Conclusion 25. Planning should begin early for constructing and maintaining a sophisticated electronic helpdesk and FAQ for DØ software installation, implementation, and use issues. Triage strategies should be planned.

Conclusion 26. It might be useful to get a non-binding expression of interest from potential institutions just to see what the maximum might be, and to determine whether interest is not sufficient to support the concept. Too few sites will make the burden on that small set perhaps too significant for their viability.

Conclusion 27. A **Prototype RAC Project** should be mounted to establish a working RAC by October 1, 2002. “Working” should be minimally defined to be a) the RAC site accepting continuous TMB files; b) and identified IAC sites in a new cluster using them with relative ease to do DØ analysis.

building Use Case for possible IAC desktop W cross section measurement



back

imagined RAC services:

enhanced batch processing for region

- authentication? firewalls?
- maybe IAC processing privileges at local RAC initially?

data cache and delivery for region

- everyone must be a SAM site
- RACs deliver not just to local IACs, but everywhere

database access for region

- hopefully can rely on db proxies

data reprocessing for collaboration

- see below for special concerns

monte carlo production, or service to MC IAC sites

- self-contained, minimal db access
- essentially done now

notice what's not there:

ab initio reconstruction

- presume farm will always keep up

code distribution

- after discussing it, there seemed to be no necessity for RACs to support code distribution outside of the currently evolving UPS/UPD based distribution started with the DØRACE workshop
- there might be a need for local support structures to triage questions/problems before they get back to FNAL
  - presumably distributed expertise with code dist., SAM, databases, etc.

Best RAC requirements, 1

location, location, location

They have to be positioned in order to serve (and protect?)

Anticipate a few RACs - not more than ~10

Didn’t try to establish firm siting criteria

- considered, but rejected, notion of analysis topic-based sites
  - *ie, not a SUSY RAC, Extra Dimensions RAC, top RAC, etc*
- rather, try to distribute according to density of users
  - *but there will be other overriding considerations:*
    - network capabilities, political issues (language, funding, national goals, etc), just plain interest, etc.

Networking capabilities

high-bandwidth, RACs to FNAL required

high-bandwidth, RAC to local IACs

nice, but not necessary, high-bandwidth, RAC to all other RACs

Best RAC requirements, 2

data storage. Tried to make a "*Best*" model

Generally thought desirable:

- all TMB files on disk at all RACs
- all DSTs on disk at the sum of all RACs -distributed randomly
  - *qualitatively different from FNAL service - complimentary*
  - *hopefully the source for most reprocessing needs*
- a variety of other formats on disk, keeping in mind MC needs may involve local, high-capacity caching
  - *rootuples or other derived formats*
  - *MC DST – depending on MC generation within cluster?*
  - *database/SAM disk storage*
  - *temporary cache ~10% of total*

results in ~100TB disk storage per RAC for Run IIa

computing. Used cpb model, guess □ 20% x fnal capability

guess ~50 nodes per year per Best RAC

Best RAC data storage

using the tools for the cpb document to the Director's Review - a model for storage:

	size	tape factor	disk factor
raw event	0.25 MB	0	0
raw/RECO	0.5 MB	0.001	0.005
data DST	0.15 MB	0.1	0.1
data TMB	0.01 MB	1	2
data root/derived	0.01 MB	0	1
MC D0Gstar	0.7 MB	0	0
MC D0Sim	0.3 MB	0	0
MC DST	0.3 MB	0.025	0.05
MC TMB	0.02 MB	0	0
PMCS MC	0.02 MB	0	0
MC rootuple	0.02 MB	0.3	0.1

for example, this means:

1 complete data set-worth of TMB on tape;

2 complete data set -worth of TMB on disk

multiples, or fractions
of the raw event count
in various formats

obviously, this is tunable

Best RAC storage, cont

Disk Storage

	1 day	1 year	phase 1 2 years	phase 2 4 years
event rate	2.16E+06	7.88E+08	1.58E+09	6.31E+09
TIER DISK data accumulation (TB)				
raw event	0.0000	0.000	0.00	0.00
raw/reprocessing	0.0054	1.971	3.94	19.71
data DST	0.0324	11.826	23.65	118.26
data TMB	0.0432	15.768	31.54	157.68
data root/derived	0.0216	7.884	15.77	78.84
MC D0Gstar	0.0000	0.000	0.00	0.00
MC D0Sim	0.0000	0.000	0.00	0.00
MC DST	0.0324	11.826	23.65	118.26
MC TMB	0.0000	0.000	0.00	0.00
PMCS MC	0.0000	0.000	0.00	0.00
MC rootuple	0.0043	1.577	3.15	15.77
cache	0.0139	5.085	10.17	50.85
db/SAM		0.500	1.00	2.00
total storage (TB)	**0.1393**	**50.852**	**102**	**509**
total storage (PB)	**0.000**	**0.051**	**0.10**	**0.51**
total storage (GB)	**139**	**50,852**	**101,704**	**508,518**

Run IIa

Run IIb

Tape Storage

	1 day	1 year	phase 1 2 years	phase 2 4 years
event rate	2.16E+06	7.88E+08	1.58E+09	6.31E+09
TAPE data accumulation (TB)				
raw event	0.5400	0.000	0.00	0.00
raw/reprocessing	0.0011	0.394	0.79	3.94
data DST	0.0324	11.826	23.65	118.26
data TMB	0.0216	7.884	15.77	78.84
data root/derived	0.0000	0.000	0.00	0.00
MC D0Gstar	0.0000	0.000	0.00	0.00
MC D0Sim	0.0000	0.000	0.00	0.00
MC DST	0.0162	5.913	11.83	59.13
MC TMB	0.0000	0.000	0.00	0.00
PMCS MC	0.0000	0.000	0.00	0.00
MC rootuple	0.0130	4.730	9.46	47.30
total storage (TB)	**0.6242**	**30.748**	**61**	**307**
total storage (PB)	**0.001**	**0.03**	**0.06**	**0.31**
total storage (GB)	**624**	**30,748**	**61,495**	**307,476**

the cpb model presumes:

25Hz rate to tape, Run IIa

50Hz rate to tape, Run IIb

events 25% larger, Run IIb

Best RAC requirements, 3

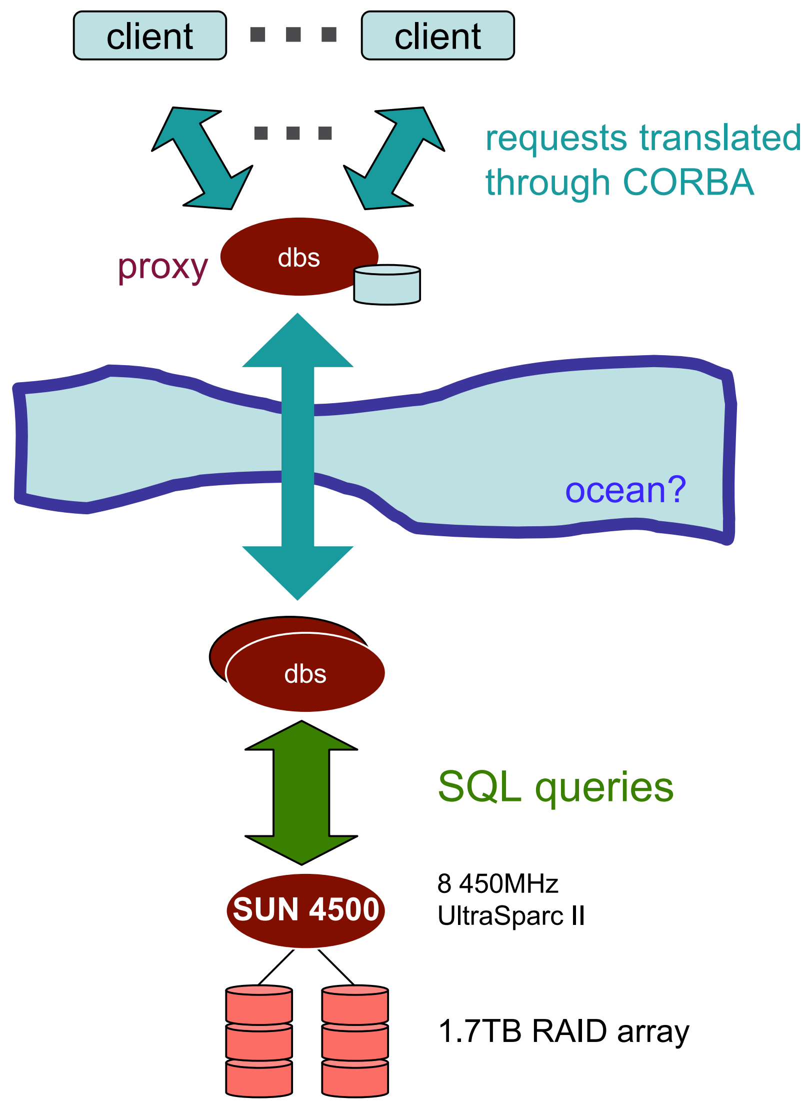
database

presume implementation of proxied database servers

- a feature of the upgrade currently under way for the server

  every RAC would house a proxy server

- this will hopefully be tested within the year



maybe some recruitment is worthwhile

imagine a Management World Tour?

so, is the argument air tight?

I think it could still use work...

Need Use Cases...which have the dual impact of:

- emphasizing how cool it could be and exposing the complications
  - The document contains a narrative for a W cross section measurement

Also need:

  - a tracking Use Case, B lifetime .
  - a high statistics Use Case, High E_T jets.
  - a reprocessing Use Case.

A we'll-fail-without-it argument

- deconstruct a few Run I analyses - what was really done - and project them onto multiple fb^{-1}
  - started to extrapolate the Run I M_W analysis
  - I think that the amount of work required for all anticipated analyses will be impossible the Old Way

Best RAC data storage

using the tools for the cpb document to the Director's Review - a model for storage:

	size	tape factor	disk factor
raw event	0.25 MB	0	0
raw/RECO	0.5 MB	0.001	0.005
data DST	0.15 MB	0.1	0.1
data TMB	0.01 MB	1	2
data root/derived	0.01 MB	0	1
MC D0Gstar	0.7 MB	0	0
MC D0Sim	0.3 MB	0	0
MC DST	0.3 MB	0.025	0.05
MC TMB	0.02 MB	0	0
PMCS MC	0.02 MB	0	0
MC rootuple	0.02 MB	0.3	0.1

multiples, or fractions
of the raw event count
in various formats

for example, this means:

1 complete data set-worth of TMB on tape;

2 complete data set -worth of TMB on disk

obviously, this is tunable

Best RAC storage, cont

Disk Storage	1 day	1 year	phase 1 2 years	phase 2 4 years
event rate	2.16E+06	7.88E+08	1.58E+09	6.31E+09
TIER DISK data accumulation (TB)				
raw event	0.0000	0.000	0.00	0.00
raw/reprocessing	0.0054	1.971	3.94	19.71
data DST	0.0324	11.826	23.65	118.26
data TMB	0.0432	15.768	31.54	157.68
data root/derived	0.0216	7.884	15.77	78.84
MC D0Gstar	0.0000	0.000	0.00	0.00
MC D0Sim	0.0000	0.000	0.00	0.00
MC DST	0.0324	11.826	23.65	118.26
MC TMB	0.0000	0.000	0.00	0.00
PMCS MC	0.0000	0.000	0.00	0.00
MC rootuple	0.0043	1.577	3.15	15.77
cache	0.0139	5.085	10.17	50.85
db/SAM		0.500	1.00	2.00
total storage (TB)	**0.1393**	**50.852**	**102**	**509**
total storage (PB)	**0.000**	**0.051**	**0.10**	**0.51**
total storage (GB)	**139**	**50,852**	**101,704**	**508,518**

Run IIa

Run IIb

Tape Storage	1 day	1 year	phase 1 2 years	phase 2 4 years
event rate	2.16E+06	7.88E+08	1.58E+09	6.31E+09
TAPE data accumulation (TB)				
raw event	0.5400	0.000	0.00	0.00
raw/reprocessing	0.0011	0.394	0.79	3.94
data DST	0.0324	11.826	23.65	118.26
data TMB	0.0216	7.884	15.77	78.84
data root/derived	0.0000	0.000	0.00	0.00
MC D0Gstar	0.0000	0.000	0.00	0.00
MC D0Sim	0.0000	0.000	0.00	0.00
MC DST	0.0162	5.913	11.83	59.13
MC TMB	0.0000	0.000	0.00	0.00
PMCS MC	0.0000	0.000	0.00	0.00
MC rootuple	0.0130	4.730	9.46	47.30
total storage (TB)	**0.6242**	**30.748**	**61**	**307**
total storage (PB)	**0.001**	**0.03**	**0.06**	**0.31**
total storage (GB)	**624**	**30,748**	**61,495**	**307,476**

the cpb model presumes:

25Hz rate to tape, Run IIa

50Hz rate to tape, Run IIb

events 25% larger, Run IIb